Towards a dynamical view of ACT-R's electrophysiological correlates

van Vugt, M. K.

Published in:
Proceedings of the 12th International Conference on Cognitive Modeling

Towards a dynamical view of ACT-R's electrophysiological correlates

Marieke K van Vugt (m.k.van.vugt@rug.nl)
Dept. of Artificial Intelligence, Nijenborgh 9
9747 AG Groningen
The Netherlands

Abstract

Cognitive neuroscience could benefit from more detailed theories about how different cognitive resources interact and how those interactions unfold over time. Cognitive architectures such as ACT-R make predictions about such interactions. Previous work has shown that the activation of ACT-R modules (i.e., cognitive resources) is correlated with oscillatory activity, primarily in the lower frequency bands. Activation of its visual, production, and retrieval modules is primarily associated with oscillations in occipital areas, while the activation of its imaginal module is primarily associated with oscillatory activity in the fronto-parietal attention network. What remained unclear in previous work was the time courses of these oscillatory correlates of ACT-R and what their role is in information processing. Here I show by examining the changes in these oscillatory correlates over time that the occipital effect associated with visual, production, and retrieval modules is primarily related to stimulus identification, and the fronto-parietal correlate associated with the imaginal module is primarily associated with target consolidation. In addition, I demonstrate how patterns of coherence between electrodes could be interpreted as information transfer between the relevant ACT-R modules.
Keywords: EEG; ACT-R; attentional blink; oscillations

Introduction

The brain is a very complex system with many degrees of freedom, which means that investigating how the brain is involved in cognition requires a strong theory. Since most model-based neuroscience (Forstmann, Wagenmakers, Eichele, Brown, & Serences, 2011) focuses on simple models of a single cognitive operation, they provide little theory about how different parts of the brain interact. In this manuscript I argue that this gap could be filled by cognitive architectures such as ACT-R, which consist of various modules (cognitive resources) that communicate and interact to produce the modeled behaviors. As such, they could provide a theory for how networks of different brain areas may together be involved in information processing.

Previous studies that related ACT-R to fMRI activity (Anderson, Fincham, Qin, & Stocco, 2008; J. Borst, Taatgen, & van Rijn, 2011) have provided a good map between specific ACT-R modules and BOLD activity and developed new insights into the neural overlap between disparate tasks (J. P. Borst & Anderson, 2013). What has been less developed are relations between ACT-R and EEG activity, even though EEG has the temporal resolution to reveal module communication in real time that fMRI lacks.

I have previously developed a mapping between ACT-R module activity and (oscillatory) EEG data during an attentional blink task (van Vugt, 2012). In that study, I used canonical correlation analysis (CCA) to find the electrodes and oscillatory frequency bands that best correlated with ACT-R's module activations. I showed that there were two sets of module activations that correlated with EEG activity. First, activation of the imaginal module was associated with parietal 4–9 Hz theta oscillations, in line with previous associations between the imaginal module and the fronto-parietal attention network (Anderson et al., 2008; J. Borst et al., 2011). Second, activation of the visual, retrieval, and production modules was mostly associated with 4–9 Hz theta activity in posterior electrodes, in line with previous associations between the visual module and occipital cortex in fMRI studies (J. Borst et al., 2011).

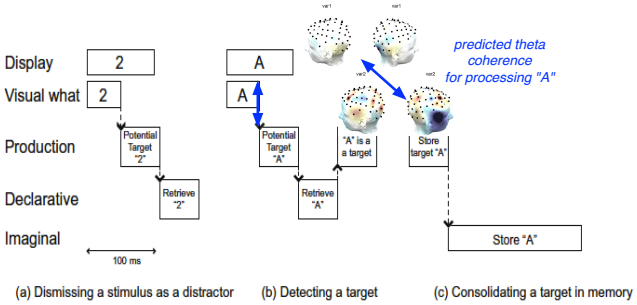

Figure 1: *Example of predictions for neural correlates of interaction between modules in the attentional blink task* (adapted from Taatgen et al., 2009). When the visual and later the production modules are identifying the letter "A," there should be a specific increase in coherence in the theta band between the channels corresponding to the stimulus identification and target consolidation variates. These sets of channels were based on Figure 3.

While this previous study introduced the approach of relating ACT-R module activation to EEG activity, it did not explore in detail how EEG-derived ACT-R activations unfolded over time. I will report here on the neural dynamics of the EEG correlates of the different ACT-R modules introduced in van Vugt (2012). An even more important aspect of ACT-R dynamics is the transfer of information between different buffers and modules, which I hypothesize may be measurable in distinct increases in coherence of the EEG (Figure 1). Coherence reflects a correlation between signals at a particular frequency, and it is thought to be a vehicle for communication in the brain (Fries, 2005). I will test whether coherence increases selectively at time points during which ACT-R predicts communication between its modules.

Methods

Task: I re-analyzed data from an attentional blink task (Martens, Munneke, Smid, & Johnson, 2006) to find the electrophysiological correlates of ACT-R. In this task, participants see a very rapid stream of visual stimuli presented for 90 ms each. Their assignment is to report whether there are letters present in the stream, and if so, which letters those are. The data reported here are from the subset of 11 blinkers in the study by Martens et al. Their EEG data were collected at the University of Groningen with a 64-channel EEG system (Twente Medical Systems, Enschede, The Netherlands) at a sample rate of 250 Hz.

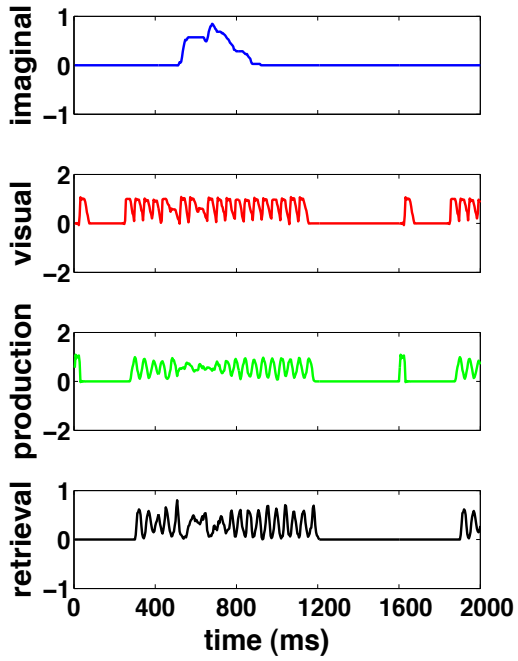

Figure 2: *Example of module activation probabilities for the ACT-R model of the attentional blink* in one of the conditions (lag 3, blinked). Blue: imaginal (working memory) module. Red: visual module. Green: production (procedural) module. Black: retrieval (from declarative memory) module. I used such module activations to create the ACT-R regressors that I correlated with the EEG data.

EEG pre-processing: I analyzed the EEG data with the EEG toolbox, a set of Matlab scripts developed in the laboratory of Michael Kahana (e.g, van Vugt, Schulze-Bonhage, Litt, Brandt, & Kahana, 2010) and custom-written scripts. I used this toolbox to extract data for every channel in our EEG setup. I concatenated the time series for each trial lengthwise into one long time series ("features") to be correlated with the ACT-R model time series ("regressors"). I then used Morlet wavelets (see van Vugt, Sederberg, & Kahana, 2007, for a comparison of different oscillation detection methods) to create representations of the EEG data in six distinct frequency bands: 2–4 Hz delta, 4–9 Hz theta, 9–14 Hz alpha, 14–28 Hz beta, 28–48 Hz low gamma and 48–90 Hz high gamma (van Vugt et al., 2010). I concatenated these frequency-transformed EEG data in the same way as the raw EEG data.

The regressors that I correlated with the EEG data reflected hypothesized model dynamics (van Vugt, Simen, & Cohen, 2011; van Vugt, Simen, Nystrom, Holmes, & Cohen, 2012). "Regressors" is fMRI terminology for a time series of interest that is used as the independent variable in a regression to find pieces of neural data that correlate with these dynamics. In this case, the data patterns of interest are probabilities of activation of the four modules that comprised the ACT-R model of the attentional blink (visual, production, retrieval, and imaginal). I ran the attentional blink ACT-R model (Taatgen, Juvina, Schipper, Borst, & Martens, 2009) 250-350 times (corresponding to the number of trials in the dataset) and computed the average activation for different model conditions: lag 3 and 8, and blinked and non-blinked responses. These average activations reflect the probability of a module being active. 'Lag' refers to the number of stimuli between the first and second target (letter) in the digit stream that the participant has to remember. An attentional blink is likely to occur for lag 3, but not lag 8 trials. Trials in which the first target was missed were removed from the analysis because in that case it is not clear what the reason is for missing the second target if that occurs.

For every trial that a participant did, I inserted the averaged module activation for the condition corresponding to that trial. This led to an activation time series during the whole task for every ACT-R module that, after subsampling to the EEG sample rate (250 Hz), had the same length as the EEG data. These were the time series that I used to regress the EEG time series on, to obtain for every module an estimate of how well it correlated with each frequency band, and which channels were most strongly involved in this correlation. Instead of using a simple multiple regression, I used a canonical correlation analysis (CCA), which finds weights on the regressors (ACT-R time series) and features (electrode time series) that maximize the correlation between the electrode- and the ACT-R time series (van Vugt et al., 2012). The CCA thus results in multiple sets of (1) a correlation value, (2) a corresponding set of weights on electrodes; the EEG "variate", and (3) a set of weights on ACT-R modules, the ACT-R "variate." Note that for this CCA, only the EEG data were wavelet-transformed–the ACT-R activation time courses were not.

Across-subject analysis: I combined the data across participants by appending the time series of the different participants in the time domain (Calhoun, Adali, Pearlson, & Pekar, 2001; Calhoun & Eichele, 2010). Similar to previous work, I used only a subset (1/4th) of each participants' data to run this CCA (van Vugt et al., 2012). This allowed me to use the remaining data in a cross-validation-based test of whether certain frequency bands showed a higher correlation with the model than others (see below). I performed

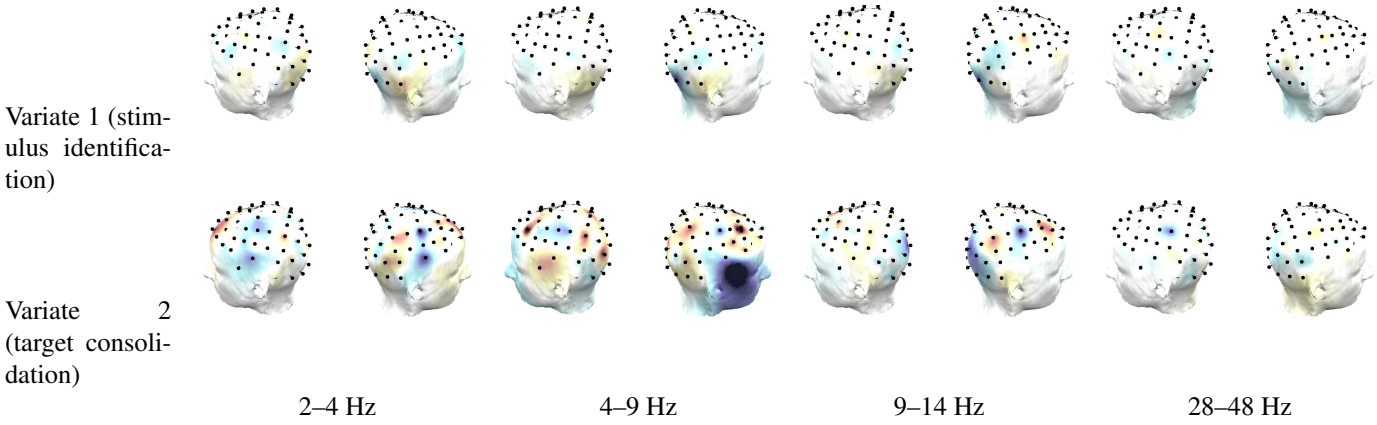

Figure 3: *Topographical representation of the correlation between ACT-R module activation and EEG.* Variate 1 primarily reflects stimulus identification done by the production and retrieval modules; variate 2 primarily reflects target consolidation done by the imaginal module. Plotted are the magnitudes of the canonical correlation weights across the brain for the canonical correlation between ACT-R time series and EEG activity in the respective frequency band. Red positive weights; blue: negative weights.

the CCA separately for every frequency band, but allowed linear combinations of regressors. This was helpful because the activations of some of the ACT-R modules were highly correlated. The CCA resulted in a set of correlations for every frequency band, one for each linear combination of regressors, and a corresponding set of electrode weights. The frequency bands with the highest correlations and the corresponding electrode weights form the EEG correlate of that particular linear combination of ACT-R modules. I assessed the significance of each canonical correlation by repeating the same analysis with a set of regressors in which the activation time courses of the respective modules were randomly displaced in time. A significant canonical correlation should exceed the distribution of canonical correlations based on 100 iterations of CCA with random data.

Comparison between frequency bands: I further tested whether the correlation in a certain frequency band was higher than that in the other bands by using a cross-validation vapproach. For every participant I applied the weights obtained with CCA to the EEG data and ACT-R time courses that had not been used in the CCA. As such, I obtained a distribution of regression coefficients (one for each participant) that I could compare between frequency bands (after having applied a Fisher transform).

Event-related averages: To compute event-related averages of the oscillatory correlates of ACT-R module activation, I applied the electrode weights to the time courses of oscillatory power and computed averages separately for each of the four trial types (lag 3 correct, lag 3 error, lag 8 correct, lag 8 error). I removed any trial that showed a kurtosis larger than 4 (Delorme & Makeig, 2004) or an average EEG amplitude larger than 70 μV. I verified by visual inspection that these thresholds removed most artifacts while at the same time retaining most non-artifactual EEG. To remove confounds of differences in overall amplitude between participants and between frequency bands, I z-scored all data before averaging.

Coherence analysis: To test the hypothesis that the different ACT-R modules communicate through coherence, I computed coherence at 7 Hz (centre of the theta band) between the sets of electrodes comprising the two canonical variates over time. I then compared coherence between the periods when there should be communication between visual and production modules (i.e., right after the activation peak of the visual module and before the peak of the production module) and when there should not be communication (i.e., during the activation peak of the visual module). This analysis was done within-subjects, and the across-subject t-statistic was computed to summarize these within-subject effects.

Results

ACT-R model of attentional blink

My analyses are based on Taatgen's previously published an ACT-R model of the the attentional blink. Figure 2 shows the time course of activation of its different ACT-R modules. The visual module (red) turns on and off each time a visual stimulus is presented (first the fixation cross; then the 20 visual stimuli that are each shown for 90 ms). The production module activates right after that to instruct the retrieval module to retrieve the identity of the stimulus, and hence its activation follows the production module activation. The imaginal module serves to store the retrieved identity into memory if the item happens to be a target. When the participant "blinks" during the trial, there is only a single activation bump of this module, indicating that the second target is never written to memory (consolidated). The two most prominent features of the attentional blink model are thus (1) module activation around each stimulus presentation and (2) a difference in imaginal module activation between blinked and non-blinked target stimuli. These should reappear in the oscillatory correlates of ACT-R.

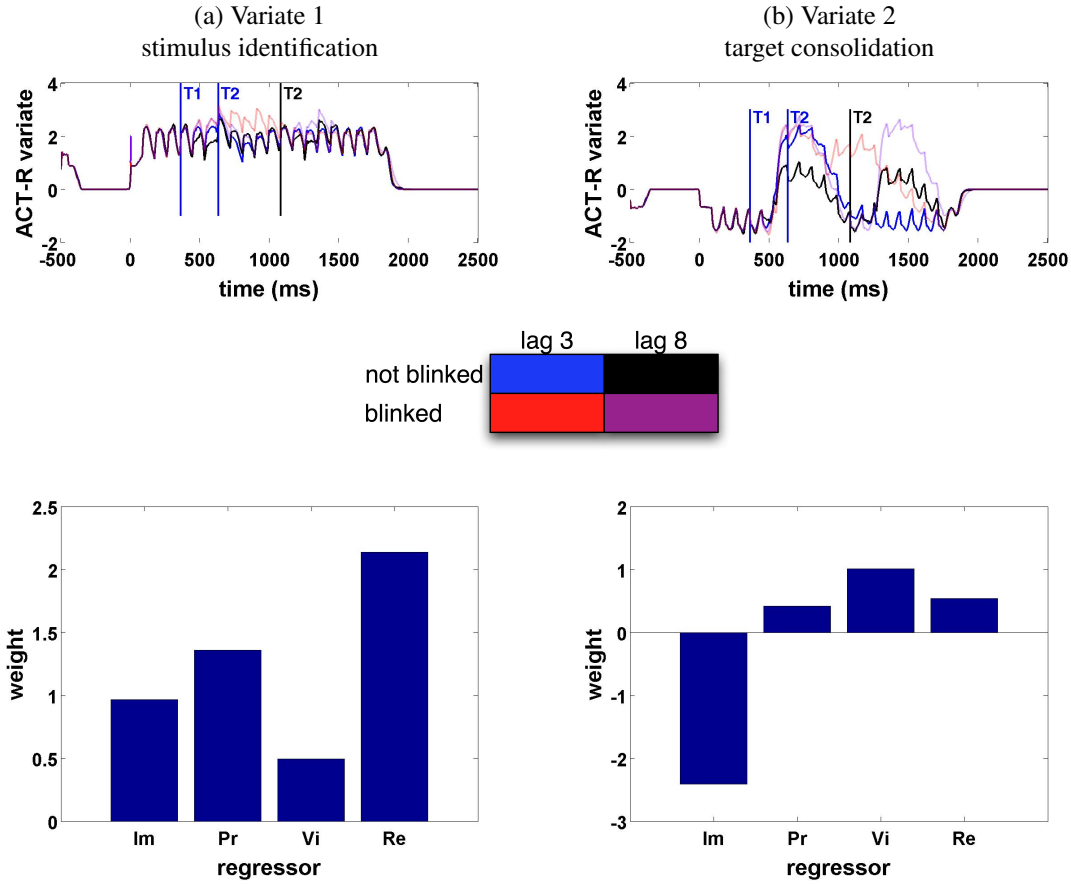

Figure 4: *Time courses of linear combinations of ACT-R module activations* that comprise variate 1 (a; on vs off-task/stimulus identification) and variate 2 (b; target consolidation) in the 4–9 Hz theta band. The four conditions are shown in different colors: lag 3 correct: blue; lag 8 correct: black; lag 3 blinked: red; lag 8 blinked: magenta. The times at which the first (T1) and second target (T2) are presented are indicated in the figure with vertical lines in corresponding colors.

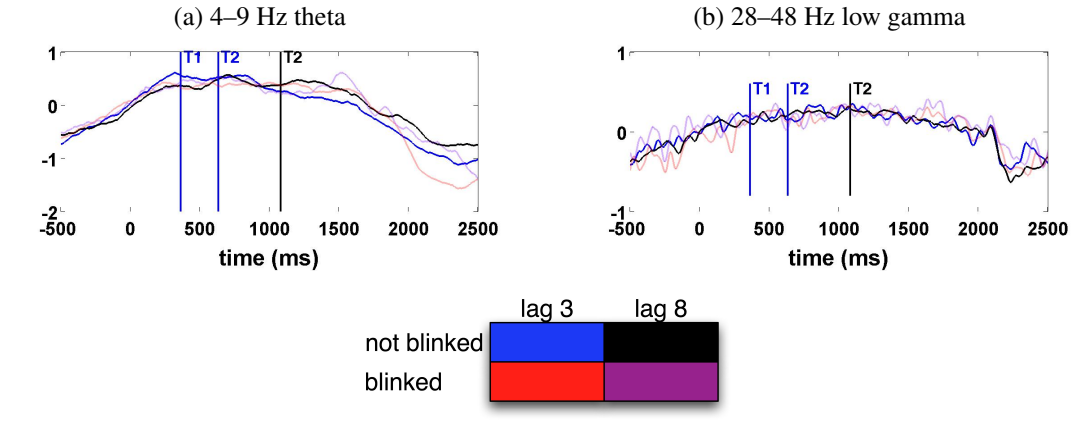

Figure 5: *Time courses of 4–9 Hz theta (a) and 28–48 Hz low gamma band activity (b) corresponding to the stimulus identification canonical variate.* The times at which the first (T1) and second target (T2) are presented are indicated in the figure with vertical lines; 0 ms is the onset of the stimulus stream. Time courses for lag 3 trials are in blue (correct) and red (blinked), while time courses for lag 8 trials are in black (correct) and magenta (blinked).

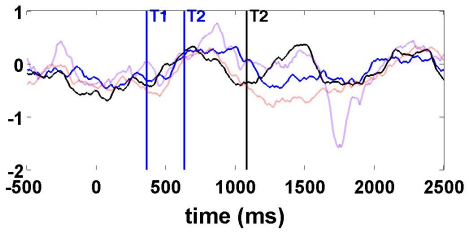

Figure 6: *Time course of 4–9 Hz theta activity corresponding to the target-consolidation canonical variate.* The times at which the first (T1) and second target (T2) are presented are indicated in the figure with vertical lines; 0 ms is the onset of the stimulus stream. Time courses for lag 3 trials are in blue (correct) and red (blinked), while time courses for lag 8 trials are in black (correct) and magenta (blinked).

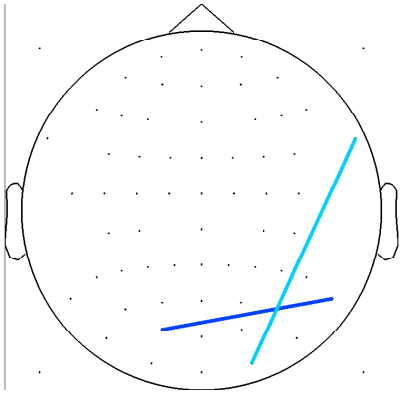

Figure 7: *Increases in coherence during model-predicted communication between visual and production modules*, which starts the process of determining the identity of the visual stimulus (what letter or digit it is).

Time course of first canonical variate

The canonical correlation analysis in van Vugt (2012) produced linear combinations of ACT-R modules that were correlated with linear combinations of EEG electrodes in different frequency bands. The first canonical variate had ACT-R weights that were largest for the production and retrieval modules. Figure 4(a) shows the weighted combination of ACT-R activations that reflects this canonical variate. There is a clear increase in amplitude for the stimulus presentation period and differentiates little between the conditions (lag 3 vs lag 8; blinked vs. non-blinked). It must therefore reflect a process that is subserved by production and retrieval modules, and that happens for every presented stimulus. A plausible candidate for such a process is identifying the identity of each visual stimulus; therefore I named this the "stimulus identification" variate.

I previously reported that this canonical variate was most strongly associated with 2–4 Hz delta and 4–9 Hz theta activity. Figure 5(a) displays the time course of 4–9 Hz theta activity in the set of channels associated with this stimulus identification variate. Does the 4–9 Hz dynamics differ from that at higher frequencies? Figure 5(b) shows the corresponding activity for the 28–48 Hz low gamma band. What is notable is that while the theta activity is mostly sensitive to being on or off-task (i.e., it turns on at the beginning of the stimulus train and turns off at its end), gamma activity also appears to be modulated by stimulus appearance (i.e., modulated by each stimulus itself).

Time course of second canonical variate

The second canonical variate loaded predominantly on the imaginal module, thus potentially reflecting the process of recognition and/or consolidation of the target stimulus. Figure 4(b) shows the weighted combination of ACT-R activations that reflects this canonical variate. For this variate, 4–9 Hz theta power does not increase and decrease with onset and offset of the stimulus stream, but rather increases more selectively after each target. For lag 3, the two target-related theta power peaks appear to be merged, while for lag 8, they can be distinguished. The second target-related peak appears to be missing for the blinked trials (red and magenta), as one would expect if this variate is involved in consolidating target stimuli, which takes place specifically for targets in non-blinked trials and is absent for targets that are blinked. I thus refer to this variate as the "target consolidation variate."

The difference in peaks between the two lags is also visible in the oscillatory correlates of this target consolidation variate shown in Figure 6. This shows that for the lag 8 condition, theta power shows 2 distinct peaks (black), while it only shows a single peak for the lag 3 condition (blue). Furthermore, theta power exhibits a narrower peak for blinked compared to non-blinked trials in the lag 3 condition (red), while it exhibits a single protracted peak for blinked trials in the lag 8 condition (magenta).

Coherence between canonical variates

Finally, I examined by way of proof of concept the idea that these correlates of ACT-R resources could be communicating through coherence (CTC hypothesis; Fries, 2005). The CTC hypothesis states that when neurons are coherent and their membrane potential goes up and down together, a neuron naturally sends most spikes at times when the other neuron is most sensitive to incoming spikes. In other words, coherence improves information exchange between neurons. Given this CTC hypothesis, I predicted that times at which ACT-R modules exchange information, there should be an increase in coherence between the channels that comprise these modules.

To test this hypothesis, I generated based on the ACT-R module time courses a set of time points at which I expected high amounts of communication, and contrasted those with time points at which there should be little communication. For example, the visual module needs to transmit information to the production module so it can ask the retrieval module about the identity of the perceived stimuli. This means there should be high coherence just after the peak activation

of the visual module and before the peak activation of the production module (Figure 1). I computed coherence for a set of electrodes in the ROIs defined by the theta topographies (Figure 3), and contrasted coherence estimates between time points of hypothesized high communication and time points of low communication. Figure 7 shows the channels that have significantly ($t(10) > 2$) increased coherence for periods where the model predicts high communication between these modules.

Discussion

In summary, I aimed to find out how the previously-identified oscillatory correlates of ACT-R behaved over the course of time. Based on visual inspection of the ACT-R components, I defined the first canonical variate observed in previous work (van Vugt, 2012) as a "stimulus identification" component, while the second canonical variate mostly reflected "target consolidation." The stimulus identification variate had previously been identified with activity in a broad range of frequencies, ranging from 2–90 Hz. Here I showed that the lower band of these frequencies primarily reflected task engagement, while the band of higher frequencies also exhibited modulation by the appearance of each stimulus. I further showed how the relation between ACT-R module activations and EEG oscillations could be used to investigate the role of oscillatory coherence in information transmission between modules. I found that communication between the visual and production modules was associated with increased coherence between occipital and parietal channels that formed part of ACT-R's oscillatory correlates.

These findings can form the basis for more understanding of the role of oscillatory coherence in information transmission in the brain. In addition, once ACT-R's electrophysiological correlates have been better established, they could also potentially be used to compare different model variants. In such model comparisons, high correlations of module activations with oscillatory power, and high correlations of information exchange between modules with coherence could provide evidence in favor of a model.

Acknowledgments

I thank all members of the University of Groningen cognitive modeling group for useful discussion on this project. Financial support was provided by Marie Curie Grant ACCDECMEM.

References

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, *12*(4), 136–143.

Borst, J., Taatgen, N., & van Rijn, H. (2011). Using a symbolic process model as input for model-based fMRI analysis: Locating the neural correlates of problem state replacements. *NeuroImage*, *58*(1), 137-147.

Borst, J. P., & Anderson, J. R. (2013). Using model-based functional MRI to locate working memory updates and declarative memory retrievals in the fronto-parietal network. *Proc. Nat. Acad. Sci., USA*, *110*(5), 1628-1633. doi: 10.1073/pnas.1221572110

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using Independent Component Analysis. *Human Brain Mapping*, *14*, 140–151.

Calhoun, V. D., & Eichele, T. (2010). Simultaneous EEG and fMRI. In M. Ullsperger & S. Debener (Eds.), (pp. 161–174). Oxford University Press. doi: 10.1093/acprof:oso/9780195372731.001.0001

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9-21.

Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., & Serences, J. T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends in Cognitive Sciences*, *15*(6), 272–279.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474-480. doi: 10.1016/j.tics.2005.08.011

Martens, S., Munneke, J., Smid, H., & Johnson, A. (2006). Quick minds don't blink: Electrophysiological correlates of individual differences in attentional selection. *Journal of Cognitive Neuroscience*, *18*(9), 1423-1438.

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive Psychology*, *59*(1), 1–29.

van Vugt, M. K. (2012). Relating ACT-R buffer activation to EEG activity during an attentional blink task. In N. Rußwinkel, U. Drewitz, & H. van Rijn (Eds.), *11th international conference on cognitive modeling - ICCM2012*. Berlin, Germany.

van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., & Kahana, M. J. (2010). Hippocampal gamma oscillations increase with working memory load. *Journal of Neuroscience*, *30*, 2694–2699.

van Vugt, M. K., Sederberg, P. B., & Kahana, M. J. (2007). Comparison of spectral analysis methods for characterizing brain oscillations. *Journal of Neuroscience Methods*, *162*(1-2), 49–63.

van Vugt, M. K., Simen, P., & Cohen, J. D. (2011). Finding neural correlates of drift diffusion processes in EEG oscillations. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (p. 2439-2444). Austin, TX: Cognitive Science Society.

van Vugt, M. K., Simen, P., Nystrom, L., Holmes, P., & Cohen, J. D. (2012). EEG oscillations reveal neural correlates of evidence accumulation. *Frontiers in Human Neuroscience*, *6*, 106.